Phys. Med. Biol. **45** (2000) 1–17. Printed in the UK PII: S0031-9155(00)12981-1

Editorial Enterprises
Marked Proof
PMB/112981/PAP
21171p
Printed on 27/09/00
at 16.00

Direct recovery of regional tracer kinetics from temporally inconsistent dynamic ECT projections using dimension-reduced time-activity basis

Jonathan S Maltz

Center for Functional Imaging, Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA

Received 30 March 2000, in final form 23 August 2000

Abstract. We present an algorithm of reduced computational cost which is able to estimate kinetic model parameters directly from dynamic ECT sinograms made up of temporally inconsistent projections. The algorithm exploits the extreme degree of parameter redundancy inherent in linear combinations of the exponential functions which represent the modes of first-order compartmental systems. The singular value decomposition is employed to find a small set of orthogonal functions, the linear combinations of which are able to accurately represent all modes within the physiologically anticipated range in a given study. The reduced-dimension basis is formed as the convolution of this orthogonal set with a measured input function. The Moore–Penrose pseudoinverse is used to find coefficients of this basis. Algorithm performance is evaluated at realistic count rates using MCAT phantom and clinical $^{99\mathrm{m}}$Tc-teboroxime myocardial study data. Phantom data are modelled as originating from a Poisson process. For estimates recovered from a single slice projection set containing 2.5×10^5 total counts, recovered tissue responses compare favourably with those obtained using more computationally intensive methods. The corresponding kinetic parameter estimates (coefficients of the new basis) exhibit negligible bias, while parameter variances are low, falling within 30% of the Cramér–Rao lower bound.

1. Introduction

Most present-day techniques for the reconstruction of emission computed tomography (ECT) images assume that the projection data are obtained from a radionuclide source distribution which does not vary in time. This is a poor assumption in most functional studies which involve the use of a rotating camera which cannot acquire projections over 360° simultaneously. There exists, consequently, a need for algorithms capable of solving the dynamic ECT reconstruction problem, which involves the estimation not only of the underlying functional anatomical source geometry, but also of the pharmacokinetics of injected radiotracer materials.

Currently, the routine method for conducting dynamic studies requires a set of reconstructed images. Each image represents the source distribution during a time interval in which activity is assumed to be stationary. Regions of interest (ROIs) are then delineated on every image. The images are then stacked in order of acquisition, and the time evolution of activity within corresponding regions across the entire image set is recorded. This process produces a time–activity curve (TAC) for each ROI. A tracer kinetic model is then fit to each TAC, using a one-dimensional regression method.

Each image of the series is produced by applying a tomographic reconstruction algorithm to a separate set of projection data, obtained during each time interval. The projection data set

0031-9155/00/010001+17$30.00 © 2000 IOP Publishing Ltd

comprises measurements taken at several angles around the distribution of radioactive sources. Implicit in the tomographic reconstruction process is the assumption that all projections are acquired from a time-invariant distribution. In dynamic studies, where the source activity is varying in time, image space methods yield biased kinetic parameters estimates so estimation must be performed directly on the measured projection data to obtain unbiased estimates.

Compartmental models are conventionally employed to describe regional tracer kinetics. For example, in myocardial single-photon-emission computed tomography (SPECT), one of the most commonly performed clinical imaging procedures in nuclear medicine, the dynamics of tracers such as $^{99\mathrm{m}}$Tc are modelled using a first-order single-compartment model. It has been shown that the wash-in parameter of this model is correlated with the perfusion of blood in the myocardium, and consequently offers a measure of myocardial viability (Smith *et al* 1994). Algorithms which have been developed for the solution of the direct estimation problem may be classified according to their assumed dynamic model. The algorithms proposed by Limber *et al* (1995) and Bauschke *et al* (1999), for example, assume that the tracer kinetics may be modelled as decaying real exponential functions. This is a special case of a first-order single-compartment model in which the blood input function is impulsive. Farnocombe *et al* (1999) extended the method of Limber *et al* by allowing recovery of all TACs that are monotonically decreasing functions of time. A limitation of these methods is that they do not allow for modelling of the wash-in phase of the tracer kinetic curves.

Reutter *et al* (1998, 1999) have demonstrated an algorithm capable of fitting single-compartment models to both phantom and clinical myocardial studies. Consequently, both tracer wash-in and wash-out parameters may be estimated. A stabilized Newton–Raphson optimization algorithm is used to solve the nonlinear weighted least squares problem whose solution yields the kinetic parameters directly from the acquired projection data. While this method is effective in providing the desired estimates, the amount of computation required is large for studies involving many dynamic regions. The objective of the approach presented here is to reduce these requirements through reduction of dimensionality and linearization of the problem.

Linear algorithms for the estimation of the kinetic parameters in dynamic ECT, which employ a preselected time–activity basis of exponential functions, have been presented in the past (Cunningham and Jones 1993, Chiao *et al* 1995). Preselection of the kinetic basis converts a problem which is nonlinear in the exponential rate parameters into a much simpler linear problem. Basis sets used by these algorithms are typically composed of families of decaying real exponential functions having rate constants selected so as to span the range of physiologically feasible modes expected in the data. For example, Cunningham and Jones (1993) utilized a set of $M = 100$ sampled exponential functions:

$$z_{\tilde{m}}[l] = \mathrm{e}^{-k_{\tilde{m}} l \Delta t} \qquad l = 0, 1, 2, \ldots, L-1 \tag{1}$$

where l is a discrete time index and $k_{\tilde{m}} \in [10^{-4},\ 1]\,\mathrm{s}^{-1}$. The $k_{\tilde{m}}$ were spaced logarithmically on this interval, whose bounds were selected for the application of exponential spectral analysis to cerebral positron emission tomography (PET) studies using three different tracer agents.

As we have shown previously, an orthogonal basis set of six functions (sampled regularly at 32 points in $t \in [0,\ 300\ \mathrm{s}]$), is able to approximate any one of these $z_{\tilde{m}}$ with a maximum deviation of well under 1% (Maltz *et al* 1998). The large dimensionality reduction which is possible illustrates the well-known high level of redundancy that exists among families of closely parametrized real decaying exponentials (Reich 1981, Lanczos 1956, p 275). Here, we exploit this redundancy to achieve significant computational savings over previous algorithms for exponential spectral analysis.

The algorithm we propose differs significantly from most of its predecessors in a philosophical sense, in that it does not endeavour to recover actual compartmental model parameters but rather the time–activity curves for each compartment. If the TACs are perfectly recovered, then it is clear that application of a one-dimensional optimization algorithm to each TAC will yield a correct set of kinetic parameter values. The solution obtained may, however, not be unique owing to poor conditioning or parameter redundancy in the model. Consequently, it seems sensible to regard the accurate extrication of the correct TACs from the projection data as the most important objective in dynamic ECT. This is the philosophy we adopt here.

2. Problem formulation

We begin by assuming that the underlying source distribution $\Omega(\boldsymbol{x})$ has been completely segmented into several regions of interest $\Omega_n(\boldsymbol{x})$, $n = 1, 2, \ldots, N$.

For the application of the algorithm to myocardial studies during which both wash-in and wash-out of the tracer occur, such as those involving $^{99\mathrm{m}}$Tc-teboroxime, we assume that tracer kinetics are governed by a single compartment model. To cope with possible region heterogeneity (O'Sullivan 1993), we incorporate additional flexibility in allowing the TAC of each ROI to be composed of linear combinations of the responses of several such models:

$$\phi_n(t) = \sum_{\tilde{m}=1}^{\tilde{M}} k_1^{\tilde{m}n} i(t) * \mathrm{e}^{-k_2^{\tilde{m}} t} \tag{2}$$

where $i(t)$ is the measured blood input function, and the '*' operator denotes convolution. The wash-in coefficient $k_1^{\tilde{m}n}$ and wash-out parameter $k_2^{\tilde{m}}$ describe the contribution of the $\tilde{m}$th mode to the time activity $\phi_n(t)$ of the nth region.

We prepare the exponential spectral basis (1) for manipulation by forming the $(L \times \tilde{M})$ matrix $\mathbf{X}$ whose $\tilde{m}$th column is $z_{\tilde{m}}[l]$ as defined in (1), and invoke the singular value decomposition (SVD) to find orthogonal basis vectors for the range of $\mathbf{X}$ (Maltz *et al* 1998). These are the left singular (column) vectors $\boldsymbol{u}_{\tilde{m}}$ of the SVD of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{U}\mathbf{S}\tilde{\mathbf{V}}^T \qquad \tilde{\mathbf{U}} = (\boldsymbol{u}_1\, \boldsymbol{u}_2 \ldots \boldsymbol{u}_{\tilde{M}}) \tag{3}$$

where $\mathbf{V}$ is the matrix of right singular vectors, and $\mathbf{S}$ is the diagonal matrix of singular values. We associate the discrete time index l with each row of $\mathbf{U}$. Depending on the degree of accuracy required in the sampled representation of the TAC's $\phi_n(t)$, we utilize only the first $M \leqslant \tilde{M}$ of $\tilde{\mathbf{U}}$ such that:

$$\mathbf{U} \equiv (\boldsymbol{u}_1\, \boldsymbol{u}_2 \ldots \boldsymbol{u}_M). \tag{4}$$

Typically, $M \approx 4$ is sufficient for the myocardial imaging applications we have studied.

We then form the matrix $\mathbf{C}$ from the columns of $\mathbf{U}$ convolved with the sampled blood input function $i[l]$, which we assume has either been measured or estimated:

$$\mathbf{C}' \equiv (\boldsymbol{c}_1\, \boldsymbol{c}_2 \ldots \boldsymbol{c}_M) \tag{5}$$

where $\boldsymbol{c}_m = \boldsymbol{u}_m * i[l]$, $l = 0, 1, \ldots, L - 1$.

With the kinetic model formalized, we wish to estimate the coefficients μ_{mn} of the $\boldsymbol{c}_m$ for all regions, which form the approximated TACs as:

$$\hat{\phi}_n[l] = \sum_{m=1}^{M} \hat{\mu}_{mn} \boldsymbol{c}_m[l] \qquad l = 0, 1, \ldots, L - 1 \tag{6}$$

where $L = RP$, the total number of angular projections, given R camera rotations with P angular projections per rotation.

3. Problem solution

Let the row vectors $\boldsymbol{y}_p^T$, of length Q each represent a single angular projection measurement, containing the measurements of Q bins along the projection. These vectors form the rows of the $RP \times Q$ measured sinogram matrix $\tilde{\mathbf{Y}}$, when P angular projections are acquired over each of R camera rotations.

We begin by forming the lexicographically stacked projection $RPQ \times 1$ vector $\tilde{\boldsymbol{y}}$ as:

$$\tilde{\boldsymbol{y}} \equiv \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_{RP} \end{bmatrix}. \tag{7}$$

We project each of the segmented regions $\Omega_n(\boldsymbol{x})$ using the same projection geometry and system response matrix that applies to the imaging system, to form the N single rotation, $P \times Q$ sinograms $\mathbf{W}_n$. We denote the P row vectors each containing an angular projection as $\boldsymbol{w}_{np}^T$, and stack the $\boldsymbol{w}_{np}$ in an identical manner to that shown in (7), giving the $PQ \times 1$ vectors $\tilde{\boldsymbol{w}}_n$.

In order that multiple rotation data may be accommodated, it is necessary to replicate the geometric weighting factors contained in the calculated sinograms R times. A similar replication is required so that we may solve for the coefficients of each of the M modes. To this end, we construct the $RPQ \times M$ matrices $\mathbf{G}_n$ which are block matrices containing $R \times M$ identical blocks of $\tilde{\boldsymbol{w}}_n$. $\mathbf{G}_n$ is given by:

$$\mathbf{G}_n \equiv \begin{array}{cc} M \text{ blocks} & \\ \begin{bmatrix} \tilde{\boldsymbol{w}}_n & \cdots & \tilde{\boldsymbol{w}}_n \\ \vdots & \vdots & \vdots \\ \tilde{\boldsymbol{w}}_n & \cdots & \tilde{\boldsymbol{w}}_n \end{bmatrix} & R \text{ blocks.} \end{array} \tag{8}$$

The geometric weighting matrix for the activity contributions of each region for all modes and all rotations is then defined as:

$$\mathbf{G} \equiv [\ \mathbf{G}_1 \quad \mathbf{G}_2 \quad \cdots \quad \mathbf{G}_N\ ]. \tag{9}$$

The second matrix we will describe consists of blocks containing the convolved basis functions $\boldsymbol{c}_m$. This matrix must be constructed so that each of the RPQ projection bin measurements, as a function of discrete time index l, may be expressed as a linear combination of the $\boldsymbol{c}_m$. For each time sample l, we form the $PQ \times M$ matrices

$$\mathbf{C}_l \equiv \begin{bmatrix} c_1[l] & c_2[l] & \cdots & c_M[l] \\ c_1[l] & c_2[l] & \cdots & c_M[l] \\ \vdots & \vdots & \vdots & \vdots \\ c_1[l] & c_2[l] & \cdots & c_M[l] \end{bmatrix} \tag{10}$$

where $c_m[l]$ is the lth element of $\boldsymbol{c}_m$, from which the $RPQ \times\ nm$ basis weighting matrix

$$\mathbf{C} \equiv \begin{bmatrix} \mathbf{C}_0 & \mathbf{C}_0 & \cdots & \mathbf{C}_0 \\ \mathbf{C}_1 & \mathbf{C}_1 & \cdots & \mathbf{C}_1 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{C}_{L-1} & \mathbf{C}_{L-1} & \cdots & \mathbf{C}_{L-1} \end{bmatrix} \tag{11}$$

is composed.

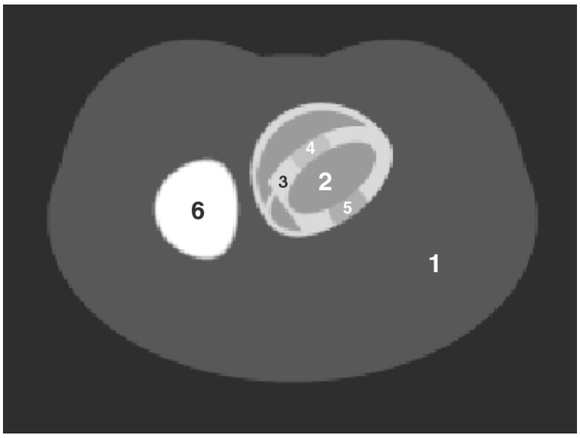

Figure 1. 2D slice of 3D MCAT emission phantom. A single slice through the myocardium transverse to the long axis of the body is taken as phantom for these studies. The liver (region 6) is shown to the left of the heart in this illustration. We see that the myocardium contains two defects (darker regions 4 and 5) and normal region 3, which is rendered non-contiguous by the defects. Region 2 is the myocardial blood pool, while region 1 represents the background activity in the torso.

The vector $\hat{\boldsymbol{\mu}}$ containing the coefficient estimates $\hat{\mu}_{mn}$ is easily obtained via solution of the linear system:

$$\tilde{\boldsymbol{y}} = \mathbf{F}\hat{\boldsymbol{\mu}} = (\mathbf{G} \cdot \mathbf{C})\hat{\boldsymbol{\mu}} \tag{12}$$

where the operator '·' denotes element-by-element multiplication.

Equation (12) may be solved directly in the least-squares sense giving:

$$\hat{\boldsymbol{\mu}} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\tilde{\boldsymbol{y}} \tag{13}$$

when $(\mathbf{F}^T\mathbf{F})$ is invertible. When this is not the case, the system of equations is ill-conditioned and the SVD may be used to find the pseudoinverse. A solution which does not depend on those singular vectors corresponding to small singular values may then be obtained. In practice, when the number of regions N is small with respect to the number of projection bin measurements, (12) is almost invariably highly overdetermined.

We henceforth refer to the algorithm developed above as the 'convolved-orthogonal basis reconstruction algorithm' (COBRA).

4. Algorithm evaluation

The algorithm is first applied to a single slice of a dynamic realistic mathematical cardiac torso (MCAT) phantom (Tsui *et al* 1993), and then to a $^{99\mathrm{m}}$Tc-teboroxime myocardial patient study.

4.1. Phantom study

The 3D MCAT phantom is shown in figure 1. This phantom models not only the myocardium, but also the myocardial blood pool, the background activity in the body, and the liver. The projections of a single slice transverse to the long axis of the body were chosen for this evaluation.

The simulated dataset was acquired over 15 rotations of a single-headed camera, taking 120 regularly spaced angular measurements per rotation, of 64 projection bins each. The data

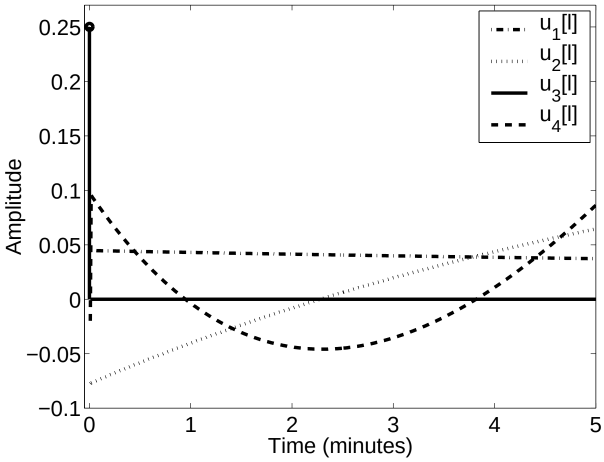

Figure 2. Orthogonal basis functions $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ and $\boldsymbol{u}_4$ together with supplemental impulsive basis function $\boldsymbol{u}_3$ are employed in the phantom studies where $M = 3$. The former are the first three vectors $\boldsymbol{u}_m$ of $\mathbf{U}$. The latter is included to model the blood pool within the imaged distribution.

were derived from a Poisson process which generated either 1×10^5, 2.5×10^5 or 1×10^6 total detected events during the 15 min imaging period. While attenuation was modelled, non-ideal system response and scatter were not.

A total of six regions, having the TACs illustrated in figure 5, were included in the phantom data.

The orthogonal basis functions were calculated through the application of the SVD to a matrix of 100 sampled decaying real exponential functions parametrized by rate constants logarithmically spaced in the interval $[5 \times 10^{-4}, 2]$. This interval includes the true range of $k_2 \in [0.002, 0.6]$ from which the TACs are derived. In practice, of course, the true range is unknown, so the choice of interval for k_2 should ensure that all physiologically feasible modes are accommodated. The number M of left singular vectors $\boldsymbol{u}_m$ retained after application of the SVD is selected as the minimum number needed to approximate all of the exponential functions $z_m[l]$ to within 1% peak deviation, using the reduced-dimension basis. An additional basis function $u_{M+1}[l] = \delta[l]$ is included to allow for explicit modelling of the blood pool within the imaged distribution, where $\delta[l]$ is the discrete-time impulse. The basis functions employed appear in figure 2.

The algorithm was tested over $I = 100$ and $I = 1000$ sinogram realizations. Owing to the parameter redundancy inherent in functions involving exponential sums, we do not attempt to recover this form of parametrization for the recovered TACs. Since the values of $\hat{\boldsymbol{\mu}}$ have no meaningful physical interpretation, and are also not guaranteed to describe each TAC uniquely, the bias and variance of parameter values themselves are considered secondary to metrics indicating the ability of the algorithm to recover TACs which are accurate estimates of the true TACs. As a measure of TAC similarity, we employ the following metric which describes the mean (over several noise realizations of the measured sinogram) of the mean RMS deviation (over all samples) of a recovered TAC $\hat{\phi}_n[l]$ as a percentage of original TAC peak value of $\phi_n[l]$, according to:

$$M_{\mathrm{dev}}^{n} \equiv \frac{1}{I} \sum_{i=1}^{I} \frac{\sqrt{\frac{1}{L} \sum_{l=1}^{L} (\phi_n[l] - \hat{\phi}_n[l])^2}}{\max_{l \in \{0,1,\ldots,L-1\}} \phi_n[l]} \times 100. \tag{14}$$

We choose to normalize by the peak original TAC value, as this allows us to conveniently compare the deviation between the two curves as a fraction of the overall scale of the true TAC. In this way, we avoid the meaningless large biases which occur in sample-by-sample comparison of two TACs, at points where the power in the true and recovered TACs is close to the noise power.

Our secondary interest is to determine the behaviour of the $\hat{\boldsymbol{\mu}}$ as a function of the data. The mean of all estimates $\bar{\hat{\boldsymbol{\mu}}}$ is compared with the true $\boldsymbol{\mu}^0$ (as recovered from noise-free projection data) to evaluate bias in the estimates. Were the projection data derived from an i.i.d. Gaussian process, we would be assured that the estimator (13) was fully efficient in that it achieves the Cramér–Rao lower bound (CRLB) on parameter variance. Since we are dealing with Poisson data, which are asymptotically Gaussian as the number of counts increases, we expect efficiency to fall off as the signal-to-noise ratio (SNR) decreases.

For those parameters which are found to be unbiased, the variance of each estimate is calculated over many noise realizations and compared with the CRLB. In order to derive the appropriate CRLB, we must first formulate the Fisher information matrix for the appropriate likelihood function. Since the data are derived from a Poisson process, their log-likelihood is given by:

$$\ln \ell(\tilde{\boldsymbol{y}}; \boldsymbol{\mu}) = \sum_{p,q} (\tilde{y}_{pq} \ln y'_{pq}(\boldsymbol{\mu}) - y'_{pq}(\boldsymbol{\mu})) + \kappa \tag{15}$$

where the measurement datum in the qth bin of the pth projection is $\tilde{y}_{pq}$, and the y'_{pq} are the elements of the vector $\boldsymbol{y}'$ obtained by evaluating the model projections (12) for a particular estimate of parameter vector $\hat{\boldsymbol{\mu}}(\tilde{\boldsymbol{y}})$ based on a particular realization of the measurement data $\tilde{\boldsymbol{y}}$:

$$\boldsymbol{y}'(\boldsymbol{\mu}) = \mathbf{F}\hat{\boldsymbol{\mu}}(\tilde{\boldsymbol{y}}). \tag{16}$$

The constant κ in (15) is independent of $\boldsymbol{\mu}$, and is eliminated in the formulation of the Fisher information matrix which is given by:

$$\mathbf{J}(\boldsymbol{\mu}) \equiv E([\nabla_{\mu}\boldsymbol{y}'(\boldsymbol{\mu}) \ln \ell(\tilde{\boldsymbol{y}}; \boldsymbol{\mu})][\nabla_{\mu}\boldsymbol{y}'(\boldsymbol{\mu}) \ln \ell(\tilde{\boldsymbol{y}}; \boldsymbol{\mu})]^T) \tag{17}$$

$$= \nabla_{\mu}\boldsymbol{y}'(\boldsymbol{\mu}) \operatorname{diag}^{-1}\{\boldsymbol{y}'(\boldsymbol{\mu})\} \nabla_{\mu}\boldsymbol{y}'(\boldsymbol{\mu})^T. \tag{18}$$

where the centre term is a diagonal matrix containing the inverse of the elements of the vector $\boldsymbol{y}'(\boldsymbol{\mu})$ (Chiao *et al* 1994).

When the estimates of the parameters $\hat{\boldsymbol{\mu}}$ are unbiased, their variance is bounded according to the Cramér–Rao inequality:

$$\mathrm{Var}[\hat{\mu}_i(\tilde{\boldsymbol{y}}) - \mu_i^0] \geqslant J^{ii} \qquad i = 0, 1, \ldots, MN \tag{19}$$

where the J^{ii} are the diagonal elements of $\mathbf{J}(\boldsymbol{\mu})^{-1}$, and μ_i and μ_i^0 are the ith elements of $\hat{\boldsymbol{\mu}}$ and $\boldsymbol{\mu}^0$ respectively (van Trees 1968, p 174). The latter vector contains the true parameters.

When the contribution of a specific basis function towards a TAC is negligible, even negligible estimation errors produce large parameter biases. Consequently, it is appropriate to perform analysis of parameter bias and variance only on those coefficients which are large enough so as to introduce significant power into the recovered TAC. To this end, we introduce the metric

$$M_{\text{pow}}^{mn} = \frac{\|\mu_{mn}\boldsymbol{c}_m\|^2}{\|\sum_{j=1}^{M}\mu_{jn}\boldsymbol{c}_j\|^2} \times 100 \tag{20}$$

where $\boldsymbol{c}_m$ is the mth convolved basis function and $\|\cdot\|$ denotes the Euclidean norm. The argument of the norm in the denominator of (20) is just the nth TAC. This metric gives the ratio of the power contained within a convolved basis function or ‘sub-TAC’ percentage of total TAC power and can hence be used to determine which coefficients should be subject to statistical analysis.

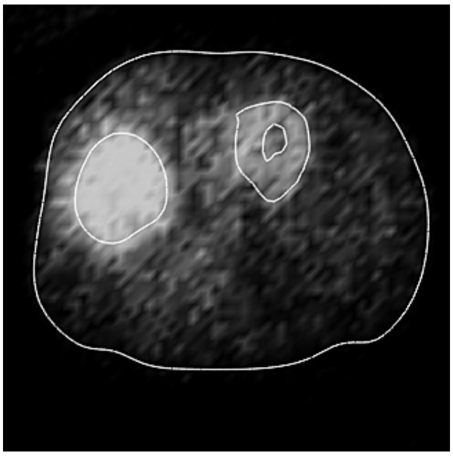

Figure 3. Specific 2D slice through imaged torso, the projections of which are selected for algorithm evaluation. The contours shown delineate tissue regions.

4.2. Patient study

To establish whether the COBRA algorithm is able to produce useful estimates of regional TAC's in a clinical setting, we apply the algorithm to a single transverse slice from a ^{99m}Tc-teboroxime myocardial patient study. While the true regional kinetics for this dataset are unknown, we are able to compare our results with those obtained previously through application to the same data of the methods of Formiconi (1993) and the direct single-compartment fit to projection data (DSCFP) algorithm of Reutter *et al* (1999).

The method of data acquisition has been described previously, but is repeated here for convenience (Reutter *et al* 1999). A dynamic SPECT study was performed at the University of Utah Medical Center, using a Picker PRISM 3000XP three-headed SPECT system. To allow for attenuation correction, an initial transmission scan was performed by mounting 65 cm focal length fan-beam collimators onto each detector head. A ^{153}Gd line source was introduced, and 120 transmission projections of 64×64 pixels were acquired over 360°. Parallel collimators were then employed for the emission study, which proceeded after vasodilation was induced by adenosine. The initiation of data acquisition coincided with the injection of 25 mCi ^{99m}Tc-teboroxime. During the total imaging time of 15 min, a full (360°) set of 120 angular projections was acquired every 10 s.

The imaged distribution was delineated into the regions: left ventricular myocardium, blood pool, liver and background tissue using the automated volume of interest specification algorithm (Reutter *et al* 1999).

The 2D slice illustrated in figure 3 was selected for the purpose of algorithm evaluation.

A primary motivation for the development of algorithms able to reconstruct imaged distributions directly from projections is the ability of such algorithms to base estimates on projection data which are temporally inconsistent. Owing to the ability of the acquisition system used in this study to perform rapid imaging, projection inconsistency does not constitute a significant problem. In order to artificially introduce projection inconsistency, we significantly decrease the time resolution of the study by summing each set of four sequentially acquired sinograms. This yields a set of 22 sinograms sampled at 40 s intervals. Since the activity of regions within the distribution changes by more than 100% during intervals of this length, a large degree of inconsistency is present in this reduced data set.

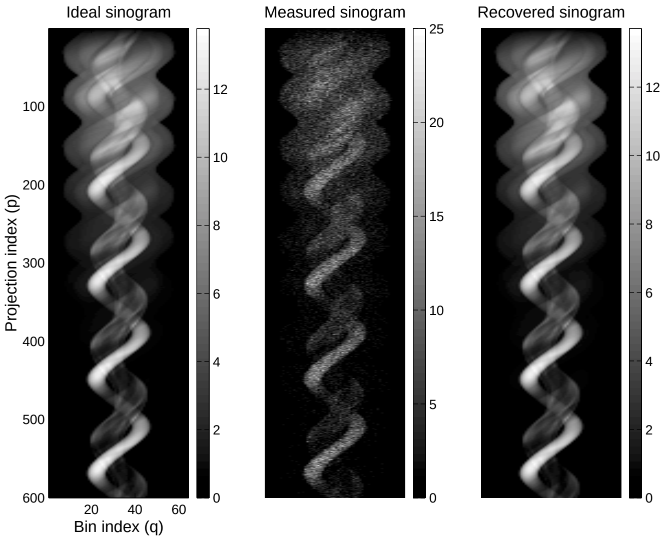
Ideal sinogram
Measured sinogram
Recovered sinogram
Projection index (p)
100
200
300
400
500
600
20
40
60
Bin index (q)
0
2
4
6
8
10
12
0
5
10
15
20
25
0
2
4
6
8
10
12

Figure 4. At 10^5 counts, the algorithm is able to reconstruct a sinogram (right) from the noisy measured sinogram (centre). The former is visually indistinguishable from the original sinogram (left).

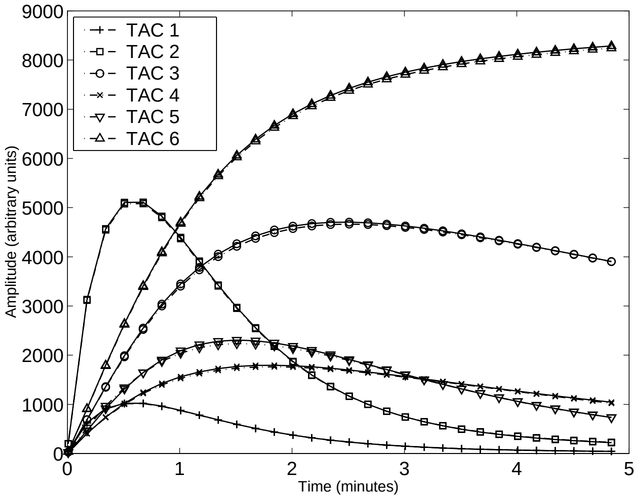

Figure 5. At 2.5×10^5 counts, the mean TAC estimates recovered over 1000 noise realizations (– –) fit the true (phantom) data very closely, illustrating a lack of discernible bias in the TACs. The original TAC samples are denoted by the symbols in the figure.

In order to compensate for attenuation, an attenuation map was constructed from truncated transmission data using 20 iterations of the ML-EM algorithm. The same algorithm was employed to reconstruct the emission data for purposes of region delineation. Registration between transmission and emission data was accomplished using three radioactive markers positioned externally around the thorax of the patient.

Table 1. Results of 100 noise realization tests of the parameter estimation algorithm.

			RMS deviation error (M_{dev}%)					
Test	Counts	Flops	TAC 1	TAC 2	TAC 3	TAC 4	TAC 5	TAC 6
1	5.0×10^5	1.1×10^8	0.18	0.58	0.75	6.98	4.35	0.27
2	2.5×10^5	1.1×10^8	0.17	0.87	1.11	8.88	5.58	0.51
3	1.0×10^5	1.1×10^8	0.19	1.09	1.54	14.69	8.44	0.46

5. Experimental results

5.1. Phantom study

Figure 4 illustrates the ability of the algorithm to estimate the kinetic parameters (coefficients of the convolved orthogonal basis set) and then reconstruct the sinogram of a dynamic distribution given the measured sinogram and a segmentation of the underlying source distribution. Figure 5 compares the original and mean recovered TACs at a total sinogram count value of 2.5×10^5, over 1000 realizations. No significant deviation between the original TACs and their estimates is discernible.

Note that we have only processed data from the first five camera rotations, as inclusion of the data obtained during the final 10 revolutions did not materially affect the estimates obtained. This behaviour stems from the highly overdetermined nature of the linear system solved by the algorithm.

Table 1 contains the results of three tests, each of 100 noise realizations at respective count totals per slice of 5×10^5, 2.5×10^5 and 1×10^5. Goodness of fit between ideal and recovered TACs is given in terms of metric M_{dev} given in (14). The average number of floating-point operations (flops) required to compute the parameters of a single sinogram noise realization also appears for each test.

Most of the errors were well below 5%, even at the lowest total counts value tested, i.e. 10^5. TAC 4, which contains the least power of all the TACs, is also the most poorly recovered, with a worst case error of $M_{\mathrm{dev}} = 14.7\%$.

Figure 6 illustrates an additional pitfall which may be encountered when a convolved orthogonal basis set is used, for which the number M of left singular vectors retained from the SVD is too large. In these results from a single noise realization test at 10^5 counts, we see that recovered TACs 4 and 5 oscillate—something which would not occur were we to use a basis of exponential functions. Oscillation in the TACs is chiefly due to the contributions of oscillatory singular vectors in the basis whose corresponding singular values are small, but whose coefficients are assigned large values by the algorithm. Intuitively, these singular vectors were ‘intended’ by the SVD to provide low-amplitude, high-frequency detail, rather than large amplitude excursions. The tremendous deterioration in performance evident in a comparison of figures 6 and 7 is due to the simultaneous increase in number of singular vectors selected for the basis (M) from 3 to 5, and drop in the total sinogram count statistic from 2.5×10^5 to 1×10^5 counts. Oscillation is avoided in figure 7 through setting $M = 3$, in effect discarding those left singular vectors whose contributions to the approximation are comparable to or smaller than the noise power. This is the rationale behind the application of the 1% peak deviation criterion on the approximation described in section 4.1.

We see from tables 2 and 3, that absolute bias is below 1% for all parameters for which M_{pow} is above 4%. Parameter 4 of TAC 4 (μ_{44}) is the most poorly estimated of all parameters, with a bias of 124%. However, the convolved basis function scaled by this parameter contains

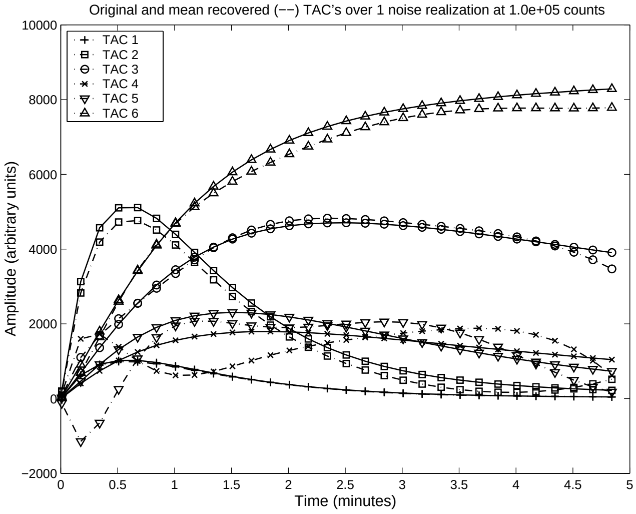

Figure 6. Oscillations in the recovered TACs may occur at low SNRs when the number of basis functions M is chosen to be too large. Here $M = 5$, whereas $M = 3$ for the results in figure 5.

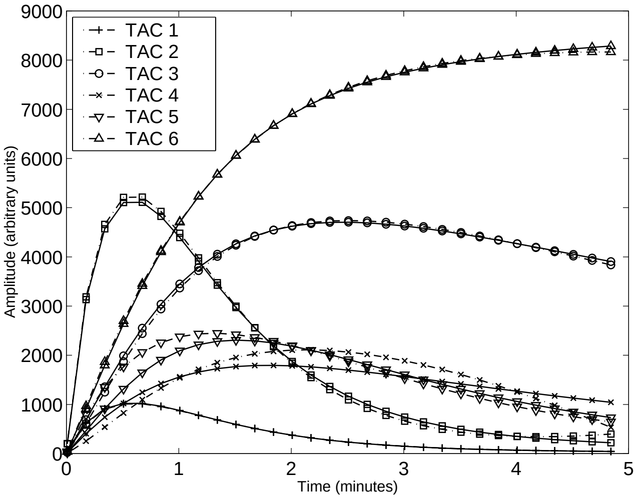

Figure 7. The oscillations present in figure 6 are not evident in the TAC estimates obtained over a typical single realization at 2.5×10^5 counts, when $M = 3$.

less than 0.1% of the total power within TAC 4, so this bias is not a significant source of error.

We see also, from table 2 and figure 8, that variances for the estimates for those coefficients which significantly weight the TACs are reasonably close to the CRLB, and do not exceed it by more than 31%. We note that this bound is only applicable to unbiased parameters, or parameters whose bias is insensitive to changes in the parameter value at the solution (van Trees 1968, p 174). Comparisons of parameter variances with their bounds may consequently be meaningless for those parameters which exhibit large bias. An example

Table 2. Quantities used in analysis of parameter bias and variance shown for TACs 1 through 3. These statistics were obtained over 1000 noise realizations, using the set of four basis functions $\boldsymbol{u}_m$ shown in figure 2 convolved with the blood input function. The 1000 measured sinograms contained 2.5×10^5 counts each over 15 revolutions. Only the first five revolutions were used to produce these results. The Cramér–Rao lower bound is abbreviated as 'CRLB'.

	TAC 1	TAC 2	TAC 3
% power in subTAC 1	42.3	42.2	95.2
% bias in coefficient 1	−0.24	−0.35	−0.61
Value	1.5×10^{-3}	3.0×10^{-4}	3.6×10^{-3}
Variance	4.4×10^{-9}	2.4×10^{-12}	4.9×10^{-9}
Var. as % of CRLB	115.56	118.05	113.46
% power in subTAC 2	55.6	55.7	2.7
% bias in coefficient 2	−0.66	−0.52	−2.66
Value	2.6×10^{-3}	-5.1×10^{-4}	-9.1×10^{-4}
Variance	1.1×10^{-8}	7.6×10^{-12}	1.2×10^{-8}
Var. as % of CRLB	115.30	112.37	
% power in subTAC 3	77.9	77.4	0.7
% bias in coefficient 3	−0.48	−0.77	0.13
Value	3.3×10^{-2}	6.6×10^{-3}	5.0×10^{-3}
Variance	2.1×10^{-6}	2.8×10^{-9}	2.6×10^{-6}
Var. as % of CRLB	114.05	120.49	111.52
% power in subTAC 4	2.6	2.5	0.0
% bias in coefficient 4	−0.21	−1.24	12.54
Value	-6.6×10^{-4}	-1.3×10^{-4}	-9.9×10^{-5}
Variance	2.8×10^{-8}	2.3×10^{-11}	3.3×10^{-8}
Var. as % of CRLB	114.51	130.09	

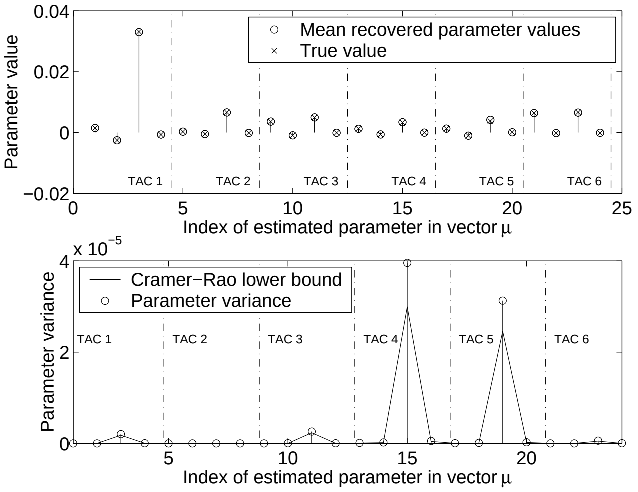

Figure 8. The upper figure shows the recovered parameter values versus their true values, and consequently indicates the bias in the recovered parameters. The lower graph compares the variance in the recovered convolved-orthogonal basis coefficients with the Cramér–Rao lower bound of each.

of such a parameter is μ_{44}. For this reason, this analysis of variance is performed only for those parameters exhibiting less than 2% absolute bias. It is interesting to note that the bias

Table 3. Quantities used in analysis of parameter bias and variance shown for TACs 4 through 6 (continuation of table 2).

	TAC 4	TAC 5	TAC 6
% power in subTAC 1	86.0	73.4	104.1
% bias in coefficient 1	-1.74	-1.40	-0.57
Value	1.2×10^{-3}	1.3×10^{-3}	6.3×10^{-3}
Variance	8.1×10^{-8}	2.6×10^{-8}	1.8×10^{-9}
Var. as % of CRLB	120.49	107.59	117.03
% power in subTAC 2	10.7	21.5	0.0
% bias in coefficient 2	-0.54	-0.77	-2.64
Value	-6.4×10^{-4}	-1.0×10^{-3}	-2.1×10^{-4}
Variance	2.0×10^{-7}	8.8×10^{-8}	4.2×10^{-9}
Var. as % of CRLB	123.88	118.42	120.81
% power in subTAC 3	2.5	3.0	0.4
% bias in coefficient 3	0.67	10.73	-0.98
Value	3.4×10^{-3}	4.2×10^{-3}	6.6×10^{-3}
Variance	4.0×10^{-5}	3.1×10^{-5}	6.0×10^{-7}
Var. as % of CRLB	132.11	127.15	120.37
% power in subTAC 4	0.0	0.1	0.0
% bias in coefficient 4	124.32	-40.58	-5.26
Value	-4.1×10^{-5}	8.3×10^{-5}	-1.1×10^{-4}
Variance	5.4×10^{-7}	2.4×10^{-7}	1.2×10^{-8}
Var. as % of CRLB	130.28	119.37	119.38

introduced through the statistical mismatch between the least squares estimator (which is optimal for i.i.d. Gaussian distributed projection bin data) and the Poisson measurement data is not appreciably deleterious. As expected, efficiency is poorer for TACs 4 and 5 which had a lower SNR.

Since the performance of this statistically suboptimal estimator appears satisfactory for clinical purposes, in that normal and compromised myocardial regions are well distinguished by their recovered TACs, we have not pursued the maximization of the Poisson likelihood function, as this requires iterative methods which are significantly more computationally intensive than the algorithm presented here, which processes this simulated data in under 15 s on a Pentium II 450 MHz CPU.

5.2. Patient study

Figure 9 compares TACs derived by applying Formiconi's method with those obtained using the COBRA algorithm presented here. Corresponding TACs appear similar, and the decreased time resolution and greater smoothness of the COBRA TACs is evident. Quantitatively, we have $M_{\mathrm{dev}} = 9.4\%$.

In figure 10, TACs derived through application of the DSCFP due to Reutter *et al* are compared with the COBRA TACs. Since the latter method was applied to the estimation of myocardial and liver activities alone, while Formiconi's method was employed to determine the background and blood pool TACs, only the two former responses are shown. Again, the two sets of curves compare favourably, with $M_{\mathrm{dev}} = 7.3\%$.

The algorithm executes in under 35 s on a Pentium II 450 MHz processor.

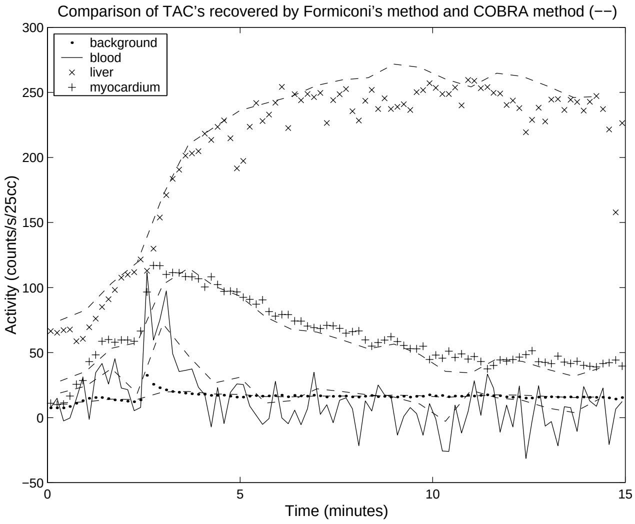

Figure 9. Comparison of TACs recovered via Formiconi's method with those obtained using the COBRA algorithm (dashed lines).

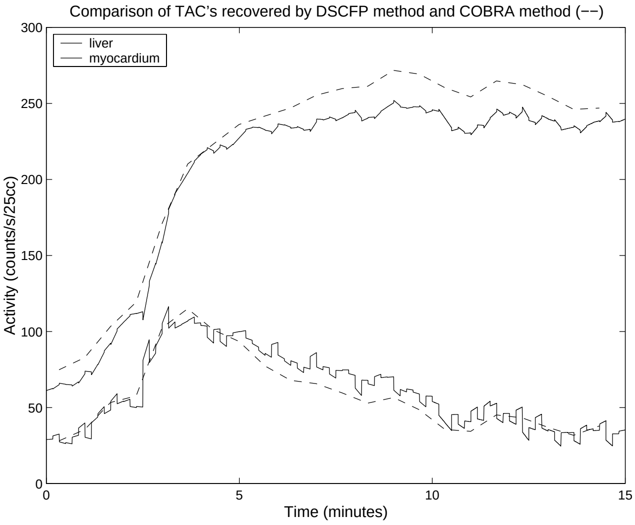

Figure 10. Comparison of TACs recovered via the direct fit of a single compartment model to projection data (method of Reutter *et al*) and the COBRA method. Blood and background TACs are not shown, as these were estimated using Formiconi's method (Reutter *et al* 1999).

6. Conclusion

The experimental results indicate that the proposed COBRA algorithm is able to rapidly recover TACs from temporally inconsistent dynamic SPECT datasets. In phantom studies, the recovered parameters typically exhibit a small bias well below 2%, and estimator efficiency is within 30% of the Cramér–Rao lower bound on parameter variance.

When applied to a clinical myocardial SPECT study rendered temporally inconsistent through artificial reduction of time-resolution, the recovered curves compared favourably with those obtained through application of the methods of Formiconi and Reutter *et al* to data with a higher temporal resolution.

The fact that good performance was obtained using a basis set of greatly reduced dimension provides empirical verification of the large degree of parameter redundancy present in the spectral formulation employed by Cunningham and Jones (1993). Consequently, there is no assurance that parameter vectors obtained using exponential spectral methods are unique, and results should be interpreted with caution. The use of the SVD, or other methods of analysis (Reich 1981), are recommended for the determination of the degree of redundancy present within the spectral basis employed in a particular study. In addition, analysis of parameter bias and variance over many noise realizations may be necessary in order to ensure that the coefficients of the spectral model cluster around a single mean parameter vector.

While many previous algorithms have required access to powerful computing equipment when applied to large multislice, multiregion studies, we have demonstrated an algorithm which scales approximately as $\mathrm{O}(RPQ(MN)^2 + (MN)^3)$ with $M \approx 4$ rather than $M \approx 100$ as in the spectral method of Cunningham and Jones The method of Reutter *et al* is more complex, scaling approximately as $\mathrm{O}(RPQN^4 + N^5)$ per iteration. However, for typical large clinical datasets with few regions (RPQ large), computation is dominated by the $\mathrm{O}(RPQN^4)$ term and the DSCFP and COBRA algorithms incur a similar computational cost for $M \approx N$. The computation time required for the application of COBRA to the clinical study using a personal computer was 35 s versus 58 s for the DSCFP algorithm.

One of the major advantages of reducing the number of kinetic parameters to be estimated is that it allows, theoretically, the estimation of the parameters of a greater number of regions at a given signal-to-noise ratio. This is important in the scenario of clinical diagnostics where wish to detect myocardial defects. Since the location of potential defects is unknown, we must segment the myocardium into several regions before estimation proceeds. This is in contrast to the clinical application of COBRA presented here, where we assumed homogeneous activity within the myocardium. The regions defined in the prior segmentation should be at least as small as the size of a physiologically significant myocardial defect. By recovering the individual TACs of the ROIs in this prior segmentation, defects may be revealed, and a diagnosis made.

The main drawback in changing the TAC basis away from the natural set of decaying real exponentials is that a significant amount of *a priori* knowledge which restricts the form of impulse response to that characteristic of first-order compartmental models is lost. At very low total count levels, where this information is important, one can expect the COBRA algorithm to yield results that are inferior to those spectral techniques which retain the exponential basis. This may be partially ameliorated by placing constraints on the magnitudes of the coefficients of those basis functions whose corresponding singular values are small. Preliminary experiments indicate that the solution of the least squares problem (13), under the constraint that all TAC estimates are non-negative is effective at preventing the recovery of physiologically infeasible TACs, even when the number of photon counts per region is decreased by an order of magnitude (Maltz 2000). Unfortunately, this requires the solution of the problem by means of an iterative technique such as quadratic programming, which increases computational requirements.

Finally, let us consider the possible limitations of the performance of the algorithm in terms of the problem-dependent parameters: the spectral range which may be accommodated, the SNR of the data and the number of regions to be resolved. The results presented here reveal no limitations in the quality of approximation using the new basis, as long as this basis is derived from only those singular vectors that are necessary to approximate the spectral range.

The recovered estimates appear robust at realistic noise levels. For studies with very low count statistics, however, the decrease in estimator efficiency observed suggests the need for an estimator better matched to the Poisson distribution of the data.

In our analysis, we have not varied the number of regions to establish how many regional TACs may be resolved at a given SNR. As the number of regions is increased, the condition of (13) will deteriorate, and some form of *a priori* knowledge will be required to restrict the solutions to those which are physiologically feasible. In preliminary studies, the imposition of simple non-negativity constraints on TAC values seems sufficient, allowing the recovery of 256- and 1024-region dynamic phantoms at respective total count values of 10^5 and 10^6, with an RMS deviation error between true and recovered TACs of approximately 20% (Maltz 2000).

Acknowledgments

The invaluable assistance of Dr B W Reutter and Dr R H Huesman in the procurement of data and preparation of this paper is much appreciated, as is the support and encouragement of Dr T F Budinger, who motivated this research. I would also like to thank Dr G T Gullberg and Dr E V R Di Bella of the Department of Radiology at the University of Utah for providing the patient data, and the Medical Imaging Research Laboratory at the University of North Carolina for making the MCAT phantom available for the simulation studies.

This work was supported in part by the National Heart, Lung, and Blood Institute of the US Department of Health and Human Services under grants HL-07367, R01-HL50663 and P01-HL25840; in part by the Director, Office of Science, Office of Biological and Environmental Research, Medical Sciences Division of the US Department of Energy under contract DE-AC03-76SF00098; and in part by the South African National Research Foundation.

References

Bauschke H H, Noll D, Celler A and Borwein J M 1999 An EM algorithm for dynamic SPECT *IEEE Trans. Med. Imaging* **18** 252–61

Chiao P, Fessler J A, Zasadny K R and Wahl R L 1995 Spectral analysis using regularized non-negative least-squares estimation *IEEE Nuclear Science Symp. and Medical Imaging Conf.* vol 3 (New York: IEEE) pp 1680–3

Chiao P C, Rogers W L, Clinthorne N H, Fessler J A and Hero A O 1994 Model-based estimation dynamic cardiac studies using ECT *IEEE Trans. Med. Imaging* **13** 217–26

Cunningham V J and Jones T 1993 Spectral analysis of dynamic PET studies *J. Cerebral Blood Flow Metabol.* **13** 15–23

Farnocombe T, Noll D, Maeght J and Harrop R 1999 An evaluation of dynamic SPECT imaging methods *IEEE Nuclear Science Symp. and Medical Imaging Conf. Record* (New York: IEEE) pp 1959–63

Formiconi A R 1993 Least squares algorithm for region-of-interest evaluation in emission tomography *IEEE Trans. Med. Imaging* **12** 90–100

Lanczos C 1956 *Applied Analysis* (Englewood Cliffs, NJ: Prentice-Hall)

Limber M A, Limber M N, Celler A, Barney J S and Borwein J M 1995 Direct reconstruction of functional parameters for dynamic SPECT *IEEE Trans. Nucl. Sci.* **42** 1249–56

Maltz J S 2000 Multiresolution constrained least-squares algorithm for direct estimation of time activity curves from dynamic ECT projection data *Proc. SPIE* **3979** 586–98

Maltz J S, Polak E and Budinger T F 1998 Multistart optimization algorithm for joint spatial and kinetic parameter estimation from dynamic ECT projection data *IEEE Nuclear Science Symp. and Medical Imaging Conf. Record* vol 3 (New York: IEEE) pp 1567–73

O'Sullivan F 1993 Imaging radiotracer model parameters in PET: a mixture analysis approach *IEEE Trans. Med. Imaging* **12** 399–411

Reich J G 1981 On parameter redundancy in curve fitting of kinetic data *Kinetic Data Analysis: Design and Analysis of Enzyme and Pharmacokinetic Experiments* ed L Endrenyi (New York: Plenum) pp 39–60

Reutter B W, Gullberg G T and Huesman R H 1998 Kinetic parameter estimation from attenuated SPECT projection measurements *IEEE Trans. Nucl. Sci.* **45** 3007–13

——1999 Kinetic parameter estimation from dynamic cardiac patient SPECT projection measurements *1998 IEEE Nuclear Science Symp. and Medical Imaging Conf. Record* pp 1953–8

Smith A M, Gullberg G T, Christian P E and Datz F L 1994 Kinetic modelling of teboroxime using dynamic SPECT imaging of a canine model *J. Nucl. Imaging* **35** 484–95

Tsui B M W, Terry J A and Gullberg G T 1993 Evaluation of cardiac cone-beam single photon emission computed tomography using observer performance experiments and receiver operating characteristic analysis *Invest. Radiol.* **28** 1101–12

van Trees H L 1968 *Detection, Estimation, and Modulation Theory* vol 1, (New York: Wiley)

Annotations from 21171p.pdf

Page 2

Annotation 1
`Farnocombe' as in reference list OK?

Annotation 2
`300 s' set as 300 seconds OK?

Page 7

Annotation 1
Please define `i.i.d.'

Page 9

Annotation 1
Sense as intended in figure 5 caption?

Page 14

Annotation 1
The key to the curves is unclear in figure 10

Page 16

Annotation 1
Publisher OK as set for `IEEE' references?